1. GRAND UNIFIED THEORIES

Updated October 2011 by S. Raby (Ohio State University).

In spite of all the successes of the Standard Model [SM] it is unlikely to be the final theory. It leaves many unanswered questions. Why the local gauge interactions $SU(3)_C \times SU(2)_L \times U(1)_Y$ and why 3 families of quarks and leptons? Moreover why does one family consist of the states $[Q, u^c, d^c; L, e^c]$ transforming as $[(3, 2, 1/3), (\bar{3}, 1, -4/3), (\bar{3}, 1, 2/3); (1, 2, -1), (1, 1, 2)]$, where $Q = (u, d)$ and $L = (\nu, e)$ are $SU(2)_L$ doublets and u^c, d^c, e^c are charge conjugate $SU(2)_L$ singlet fields with the $U(1)_Y$ quantum numbers given? [We use the convention that electric charge $Q_{EM} = T_{3L} + Y/2$ and all fields are left handed.] Note the SM gauge interactions of quarks and leptons are completely fixed by their gauge charges. Thus, if we understood the origin of this charge quantization, we would also understand why there are no fractionally charged hadrons. Finally, what is the origin of quark and lepton masses; the family mass hierarchy and quark mixing angles? Perhaps if we understood this, we would also understand neutrino masses, the origin of CP violation, the cosmological matter - antimatter asymmetry or even the nature of dark matter.

In the Standard Model, quarks and leptons are on an equal footing; both fundamental particles without substructure. It is now clear that they may be two faces of the same coin; unified, for example, by extending QCD (or $SU(3)_C$) to include leptons as the fourth color, $SU(4)_C$. The complete Pati-Salam gauge group is $SU(4)_C \times SU(2)_L \times SU(2)_R$ with the states of one family $[(Q, L), (Q^c, L^c)]$ transforming as $[(4, 2, 1), (\bar{4}, 1, \bar{2})]$ where $Q^c = (d^c, u^c)$, $L^c = (e^c, \nu^c)$ are doublets under $SU(2)_R$. Electric charge is now given by the relation $Q_{EM} = T_{3L} + T_{3R} + 1/2(B - L)$ and $SU(4)_C$ contains the subgroup $SU(3)_C \times (B - L)$ where B (L) is baryon (lepton) number. Note ν^c has no SM quantum numbers and is thus completely "sterile." It is introduced to complete the $SU(2)_R$ lepton doublet. This additional state is desirable when considering neutrino masses.

Although quarks and leptons are unified with the states of one family forming two irreducible representations of the gauge group; there are still 3 independent gauge couplings (two if one also imposes parity, *i.e.* $L \leftrightarrow R$ symmetry). As a result the three low energy gauge couplings are still independent arbitrary parameters. This difficulty is resolved by embedding the SM gauge group into the simple unified gauge group, Georgi-Glashow $SU(5)$, with one universal gauge coupling α_G defined at the grand unification scale M_G. Quarks and leptons still sit in two irreducible representations, as before, with a **10** = $[Q, u^c, e^c]$ and **$\bar{5}$** = $[d^c, L]$. Nevertheless, the three low energy gauge couplings are now determined in terms of two independent parameters : α_G and M_G. Hence, there is one prediction.

In order to break the electroweak symmetry at the weak scale and give mass to quarks and leptons, Higgs doublets are needed which can sit in either a **$5_{\mathbf{H}}$** or **$\bar{5}_{\mathbf{H}}$**. The additional 3 states are color triplet Higgs scalars. The couplings of these color triplets violate baryon and lepton number and nucleons decay via the exchange of a single color triplet Higgs scalar. Hence, in order not to violently disagree with the non-observation of nucleon decay, their mass must be greater than $\sim 10^{10-11}$ GeV. Note, in supersymmetric GUTs, in order to

cancel anomalies as well as give mass to both up and down quarks, both Higgs multiplets $\mathbf{5_H}$, $\mathbf{\bar{5}_H}$ are required. As we shall discuss later, nucleon decay now constrains the color triplet Higgs states in a SUSY GUT to have mass significantly greater than M_G.

Complete unification is possible with the symmetry group $SO(10)$ with one universal gauge coupling α_G and one family of quarks and leptons sitting in the 16 dimensional spinor representation $\mathbf{16} = [\mathbf{10} + \mathbf{\bar{5}} + \mathbf{1}]$. The SU(5) singlet $\mathbf{1}$ is identified with ν^c. $SO(10)$ has two inequivalent maximal breaking patterns. $SO(10) \rightarrow \text{SU}(5) \times \text{U}(1)_X$ and $SO(10) \rightarrow \text{SU}(4)_C \times \text{SU}(2)_L \times \text{SU}(2)_R$. In the first case we obtain Georgi-Glashow SU(5) if Q_{EM} is given in terms of SU(5) generators alone or so-called flipped SU(5) if Q_{EM} is partly in U(1)$_X$. In the latter case we have the Pati-Salam symmetry. If $SO(10)$ breaks directly to the SM at M_G, then we retain the prediction for gauge coupling unification. However more possibilities for breaking (hence, more breaking scales and more parameters) are available in $SO(10)$. Nevertheless, with one breaking pattern $SO(10) \rightarrow \text{SU}(5) \rightarrow \text{SM}$, where the last breaking scale is M_G, the predictions from gauge coupling unification are preserved. The Higgs multiplets in minimal $SO(10)$ are contained in the fundamental $\mathbf{10_H} = [\mathbf{5_H}, \ \mathbf{\bar{5}_H}]$ representation. Note only in $SO(10)$ does the gauge symmetry distinguish quark and lepton multiplets from Higgs multiplets. Finally, larger symmetry groups have been considered, *e.g.* $E(6)$, SU(6), *etc.* They however always include extra, unwanted states; making these larger symmetry groups unattractive starting points for model building.

Let us now consider the primary GUT prediction, i.e gauge coupling unification. The GUT symmetry is spontaneously broken at the scale M_G and all particles not in the SM obtain mass of order M_G. When calculating Green's functions with external energies $E \gg M_G$, we can neglect the mass of all particles in the loop and hence, all particles contribute to the renormalization group running of the universal gauge coupling. However, for $E \ll M_G$ one can consider an effective field theory (EFT) including only the states with mass $< E \ll M_G$. The gauge symmetry of the EFT [valid below M_G] is SU(3)$_C \times$ SU(2)$_L \times$ U(1)$_Y$ and the three gauge couplings renormalize independently. The states of the EFT include only those of the SM; 12 gauge bosons, 3 families of quarks and leptons and one or more Higgs doublets. At M_G the two effective theories [the GUT itself is most likely the EFT of a more fundamental theory defined at a higher scale] must give identical results; hence we have the boundary conditions $g_3 = g_2 = g_1 \equiv g_G$ where at any scale $\mu < M_G$ we have $g_2 \equiv g$ and $g_1 = \sqrt{5/3} \ g'$. [Note, the hypercharge coupling is rescaled in order for Y to satisfy the charge quantization of the GUT. Also $\alpha_s = (g_3^2/4\pi)$, $\alpha_{EM} = (e^2/4\pi)$ ($e = g \sin\theta_W$) and $\sin^2\theta_W = (g')^2/(g^2 + (g')^2)$.] Then using two low energy couplings, such as $\alpha_s(M_Z)$, $\alpha_{EM}(M_Z)$, the two independent parameters α_G, M_G can be fixed. The third gauge coupling, $\sin^2\theta_W$ in this case, is then predicted. This was the procedure up until about 1991. Subsequently, the uncertainties in $\sin^2\theta_W$ were reduced ten fold. Since then, $\alpha_{EM}(M_Z)$, $\sin^2\theta_W$ have been used as input to predict α_G, M_G and $\alpha_s(M_Z)$.

Note, the above boundary condition is only valid when using one loop renormalization group [RG] running. With precision electroweak

data, however, it is necessary to use two loop RG running. Hence, one must include one loop threshold corrections to gauge coupling boundary conditions at both the weak and GUT scales. In this case it is always possible to define the GUT scale as the point where $\alpha_1(M_G) = \alpha_2(M_G) \equiv \tilde{\alpha}_G$ and $\alpha_3(M_G) = \tilde{\alpha}_G (1 + \epsilon_3)$. The threshold correction ϵ_3 is a logarithmic function of all states with mass of order M_G and $\tilde{\alpha}_G = \alpha_G + \Delta$ where α_G is the GUT coupling constant above M_G and Δ is a one loop threshold correction. To the extent that gauge coupling unification is perturbative, the GUT threshold corrections are small and calculable. This presumes that the GUT scale is sufficiently below the Planck scale or any other strong coupling extension of the GUT, such as a strongly coupled string theory.

Supersymmetric grand unified theories [SUSY GUTs] are an extension of non-SUSY GUTs. The key difference between SUSY GUTs and non-SUSY GUTs is the low energy effective theory which, in a SUSY GUT, also satisfies N=1 supersymmetry down to scales of order the weak scale. Hence, the spectrum includes all the SM states plus their supersymmetric partners. It also includes one pair (or more) of Higgs doublets; one to give mass to up-type quarks and the other to down-type quarks and charged leptons. Two doublets with opposite hypercharge Y are also needed to cancel fermionic triangle anomalies. Note, a low energy SUSY breaking scale (the scale at which the SUSY partners of SM particles obtain mass) is necessary to solve the gauge hierarchy problem.

Simple non-SUSY SU(5) is ruled out; initially by the increased accuracy in the measurement of $\sin^2\theta_W$ and by early bounds on the proton lifetime (see below). However, by now LEP data has conclusively shown that SUSY GUTs is the new standard model; by which we mean the theory used to guide the search for new physics beyond the present SM. SUSY extensions of the SM have the property that their effects decouple as the effective SUSY breaking scale is increased. Any theory beyond the SM must have this property simply because the SM works so well. However, the SUSY breaking scale cannot be increased with impunity, since this would reintroduce a gauge hierarchy problem. Unfortunately there is no clear-cut answer to the question, when is the SUSY breaking scale too high. A conservative bound would suggest that the third generation quarks and leptons must be lighter than about 1 TeV, in order that the one loop corrections to the Higgs mass from Yukawa interactions remains of order the Higgs mass bound itself.

At present gauge coupling unification within SUSY GUTs works extremely well. Exact unification at M_G, with two loop renormalization group running from M_G to M_Z, and one loop threshold corrections at the weak scale, fits to within 3 σ of the present precise low energy data. A small threshold correction at M_G ($\epsilon_3 \sim -4\%$) is sufficient to fit the low energy data precisely.* This may be compared to non-SUSY GUTs where the fit misses by $\sim 12\,\sigma$

---

* This result implicitly assumes universal GUT boundary conditions for soft SUSY breaking parameters at M_G. In the simplest case we have a universal gaugino mass $M_{1/2}$, a universal mass for squarks and sleptons m_{16} and a universal Higgs mass m_{10}, as motivated by $SO(10)$. In some cases, threshold corrections to gauge coupling unification can be exchanged for threshold corrections to soft SUSY parameters.

and a precise fit requires new weak scale states in incomplete GUT multiplets or multiple GUT breaking scales.

Baryon number is necessarily violated in any GUT. In SU(5) nucleons decay via the exchange of gauge bosons with GUT scale masses, resulting in dimension 6 baryon number violating operators suppressed by $(1/M_G^2)$. The nucleon lifetime is calculable and given by $\tau_N \propto M_G^4/(\alpha_G^2 \, m_p^5)$. The dominant decay mode of the proton (and the baryon violating decay mode of the neutron), via gauge exchange, is $p \to e^+ \, \pi^0 \, (n \to e^+ \, \pi^-)$. In any simple gauge symmetry, with one universal GUT coupling and scale $(\alpha_G, \, M_G)$, the nucleon lifetime from gauge exchange is calculable. Hence, the GUT scale may be directly observed via the extremely rare decay of the nucleon. Experimental searches for nucleon decay began with the Kolar Gold Mine, Homestake, Soudan, NUSEX, Frejus, HPW, and IMB detectors. The present experimental bounds come from Super-Kamiokande and Soudan II. We discuss these results shortly. Non-SUSY GUTs are also ruled out by the non-observation of nucleon decay. In SUSY GUTs, the GUT scale is of order 3×10^{16} GeV, as compared to the GUT scale in non-SUSY GUTs which is of order 10^{15} GeV. Hence, the dimension 6 baryon violating operators are significantly suppressed in SUSY GUTs with $\tau_p \sim 10^{34-38}$ yrs.

However, in SUSY GUTs there are additional sources for baryon number violation – dimension 4 and 5 operators. Although the notation does not change, when discussing SUSY GUTs all fields are implicitly bosonic superfields and the operators considered are the so-called F terms which contain two fermionic components and the rest scalars or products of scalars. Within the context of SU(5) the dimension 4 and 5 operators have the form $(\mathbf{10} \, \bar{\mathbf{5}} \, \bar{\mathbf{5}}) \supset (u^c \, d^c \, d^c) + (Q \, L \, d^c) + (e^c \, L \, L)$ and $(\mathbf{10} \, \mathbf{10} \, \mathbf{10} \, \bar{\mathbf{5}})$ $\supset (Q \, Q \, Q \, L) + (u^c \, u^c \, d^c \, e^c) + B$ and L conserving terms, respectively. The dimension 4 operators are renormalizable with dimensionless couplings; similar to Yukawa couplings. On the other hand, the dimension 5 operators have a dimensionful coupling of order $(1/M_G)$.

The dimension 4 operators violate baryon number or lepton number, respectively, but not both. The nucleon lifetime is extremely short if both types of dimension 4 operators are present in the low energy theory. However both types can be eliminated by requiring R parity. In SU(5) the Higgs doublets reside in a $\mathbf{5_H}$, $\bar{\mathbf{5}}_\mathbf{H}$ and R parity distinguishes the $\bar{\mathbf{5}}$ (quarks and leptons) from $\bar{\mathbf{5}}_\mathbf{H}$ (Higgs). R parity (or its cousin, family reflection symmetry takes $F \to -F$, $H \to H$ with $F = \{\mathbf{10}, \, \bar{\mathbf{5}}\}$, $H = \{\bar{\mathbf{5}}_\mathbf{H}, \, \mathbf{5_H}\}$. This forbids the dimension 4 operator $(\mathbf{10} \, \bar{\mathbf{5}} \, \bar{\mathbf{5}})$, but allows the Yukawa couplings of the form $(\mathbf{10} \, \bar{\mathbf{5}} \, \bar{\mathbf{5}}_\mathbf{H})$ and $(\mathbf{10} \, \mathbf{10} \, \mathbf{5_H})$. It also forbids the dimension 3, lepton number violating, operator $(\bar{\mathbf{5}} \, \mathbf{5_H}) \supset (L \, H_u)$ with a coefficient with dimensions of mass which, like the μ parameter, could be of order the weak scale and the dimension 5, baryon number violating, operator $(\mathbf{10} \, \mathbf{10} \, \mathbf{10} \, \bar{\mathbf{5}}_\mathbf{H})$ $\supset (Q \, Q \, Q \, H_d) + \cdots$. Note, R parity is the only known symmetry [consistent with a SUSY GUT] which can prevent unwanted dimension four operators. Hence, by naturalness arguments, R parity must be a symmetry in the effective low energy theory of any SUSY GUT. This does not mean to say that R parity is guaranteed to be satisfied in any GUT.

Dimension 5 baryon number violating operators are generically generated via color triplet Higgsino exchange. Hence, the color triplet partners of Higgs doublets must necessarily obtain mass of order the GUT scale. The dominant decay modes from dimension 5 operators are $p \to K^+ \bar{\nu}$ $(n \to K^0 \bar{\nu})$. Note, final states with a second or third generation particle are dominant. This is due to a simple symmetry argument. The operators $(Q_i Q_j Q_k L_l)$, $(u_i^c u_j^c d_k^c e_l^c)$ (where i, j, k, $l = 1, 2, 3$ are family indices and color and weak indices are implicit) must be invariant under $SU(3)_C$ and $SU(2)_L$. Hence, their color and weak doublet indices must be anti-symmetrized. However this product of bosonic superfields must be totally symmetric under interchange of all indices. Thus, the first operator vanishes for $i = j = k$ and the second vanishes for $i = j$.

Recent Super-Kamiokande bounds on the proton lifetime severely constrain these dimension 6 and 5 operators with $\tau_{(p \to e^+ \pi^0)} > 5.0 \times 10^{33}$ yrs (79.3 ktyr exposure), $\tau_{(n \to e^+ \pi^-)} > 5 \times 10^{33}$ yrs (61 ktyr), and $\tau_{(p \to K^+ \bar{\nu})} > 1.6 \times 10^{33}$ yrs (79.3 ktyr), $\tau_{(n \to K^0 \bar{\nu})} > 1.7 \times 10^{32}$ yrs (61 ktyr) at (90% CL) based on the listed exposures. These constraints are now sufficient to rule out minimal SUSY SU(5). However non-minimal Higgs sectors in SU(5) or minimal $SO(10)$ theories still survive. The upper bound on the proton lifetime from these theories are approximately a factor of 5 above the experimental bounds. Hence, if SUSY GUTs are correct, nucleon decay must be seen soon.

Is there a way out of this conclusion? String theories, and recent field theoretic constructions, contain grand unified symmetries realized in higher dimensions. Upon compactification to four dimensions, the GUT symmetry is typically broken directly to the MSSM. A positive feature of this approach is that the color triplet Higgs states are projected out of the low energy spectrum. At the same time, quark and lepton states now emanate from different GUT multiplets. As a consequence, proton decay due to dimension 5 and 6 operators can be severely suppressed, eliminated all together or sometimes even enhanced. Hence, the observation of proton decay may distinguish extra-dimensional GUTs from four dimensional ones.

Grand unification of the strong and electroweak interactions at a unique high energy scale $M_G \sim 3 \times 10^{16}$ GeV requires [1] gauge coupling unification, [2] low energy supersymmetry [with a large SUSY desert], and [3] nucleon decay. The first prediction has already been verified. Perhaps the next two will be seen soon. Whether or not Yukawa couplings unify is more model dependent. Nevertheless, the 16 dimensional representation of quarks and leptons in $SO(10)$ is very compelling and may yet lead to an understanding of fermion masses and mixing angles. GUTs also make predictions for Yukawa coupling unification, they provide a natural framework for neutrino masses and mixing angles, magnetic monopoles, baryogenesis, *etc*. For a more comprehensive discussion of GUTs, see the unabridged particle data book. In any event, the experimental verification of the first three pillars of SUSY GUTs would forever change our view of Nature. Moreover, the concomitant evidence for a vast SUSY desert would expose a huge lever arm for discovery. For then it would become clear that experiments probing the TeV scale could reveal physics at the GUT scale and perhaps beyond.

Further discussion and references may be found in the full *Review*.